August 11th 2011 – OSG Site Admin Workshop/CMS Tier3 Meeting

Jason Zurawski, Internet2 Research Liaison

Networking Initiatives for Science

Agenda

- Internet2 Support of Science & Research
- LHC Computing Changes – R&E Networking Response
- Recent Initiatives of Interest
  - NDDI & OS^3E
  - LHCONE
  - DYNES
- Conclusions

Internet2 Seven Focus Areas

- Advanced network and network services leadership
- Services at scale: Services “above the network”
- U.S. UCAN
- National/Regional collaboration
- Global reach and leadership
- Research community development and engagement
- Industry partnership development and engagement

Research Community Engagement

- Advanced Network
  - Upgrades to 100G in the coming years
- Open Source Software
  - Monitoring, Traffic Management, Identity Management
- Services
  - Workshops, Performance Debugging and Validation, Dynamic Network Capacity, Certificate Management

Scientific Networking Needs

- Support for Science:
  - Science and engineering communities are using high-performance networking for:
    - Interactive collaboration
    - Distributed data storage and data mining
    - Large-scale, multi-site computation
    - Real-time access to remote resources
    - Dynamic data visualization
    - Shared Virtual Reality

INTERNET2

Scientific Networking Needs

- Communities Served:
  - Physics (High/Medium Energy, Nuclear)
  - Astronomy
  - Climate and Weather
  - Marine Science/Oceanography
  - Biology
  - Energy Research
  - Etc. etc. etc.
- Beyond Science
  - Arts and Humanities
  - K-20
  - Medicine*
  - First Responders (e.g. emergency services)*

* New, via the U.S. UCANN Work – http://usucann.org

Scientific Networking Needs

- Data movement to support science:
  - Increasing in size (100s of TBs in the LHC World)
  - Becoming more frequent (multiple times per day)
  - Reaching more consumers (VO size stands to increase)
  - Time sensitivity (data may grow "stale" if not processed immediately)

Scientific Networking Needs

- Traditional networking:
  - R&E or Commodity "IP" connectivity is subject to use by other users
  - Supporting large sporadic flows is challenging for the engineers, and frustrating for the scientists
  - "Speed "

INTERNET2

“Horror Stories”

- Not all communities are as prepared as others:
  - Genomics Example
    - Genome Sequencer price is falling
    - Storage and compute are cheap, cloud computing is cheaper …
    - What does this do to the networking requirements of a small school?
  - “Sneaker net”
    - Some communities may still rely on FedEx or a USB stick to move research
    - Worked well, as long as the research was <= size of portable storage
    - Starting to exceed this – lots of new demand coming …

Agenda

- Internet2 Support of Science & Research
- LHC Computing Changes – R&E Networking Response
- Recent Initiatives of Interest
  - NDDI & OS^{3}E
  - LHCONE
  - DYNES
- Conclusions

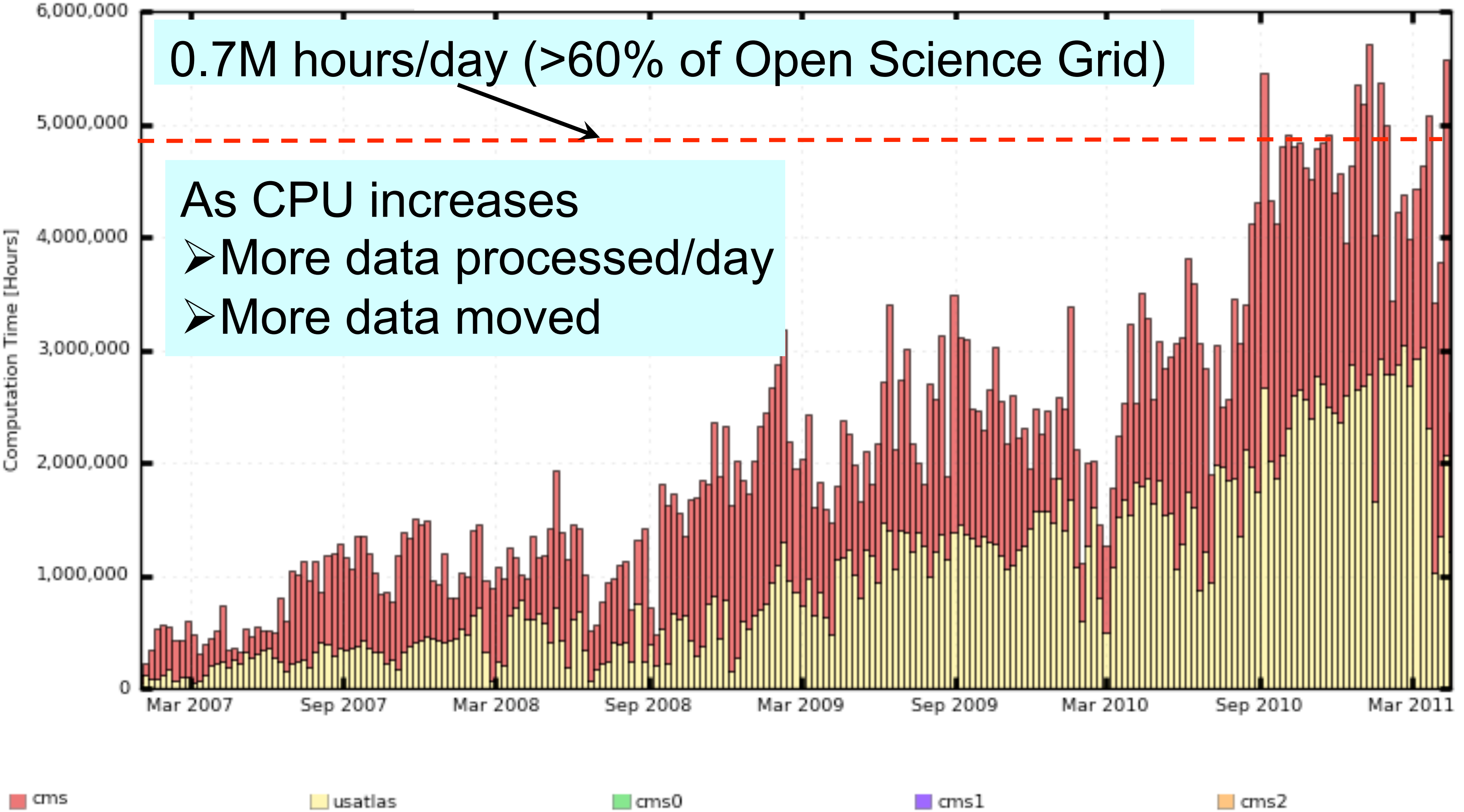
Our Relationship with LHC

- LHC communities are a model to follow:
  - Well stated requirements
  - Capacity planning years in advance
  - Tools capable of fully utilizing the network
- Prior Work
  - LHCOPN
  - USATLAS (T1s and T2s)
  - REDDnet & Vanderbilt University
  - Emerging work with CMS (T2s and T3s in the US)

Computing Growth: US-ATLAS & US-CMS

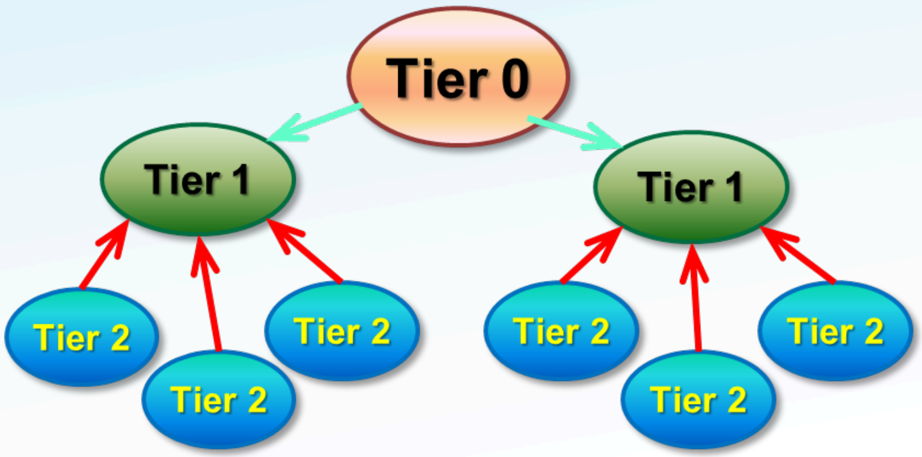
Source: P. Avery

Jan. 2007 – Apr. 2011



LHC Traditional Model

- MONARC
  - Reliance on Tiers
  - Move data between affiliated sites
  - Requirements (e.g. flow sizes, durations) are still well stated
- Lots of practice
  - In use for several years
- Dedicated Networking
  - LHCOPN
  - Private connectivity in the lower tiers

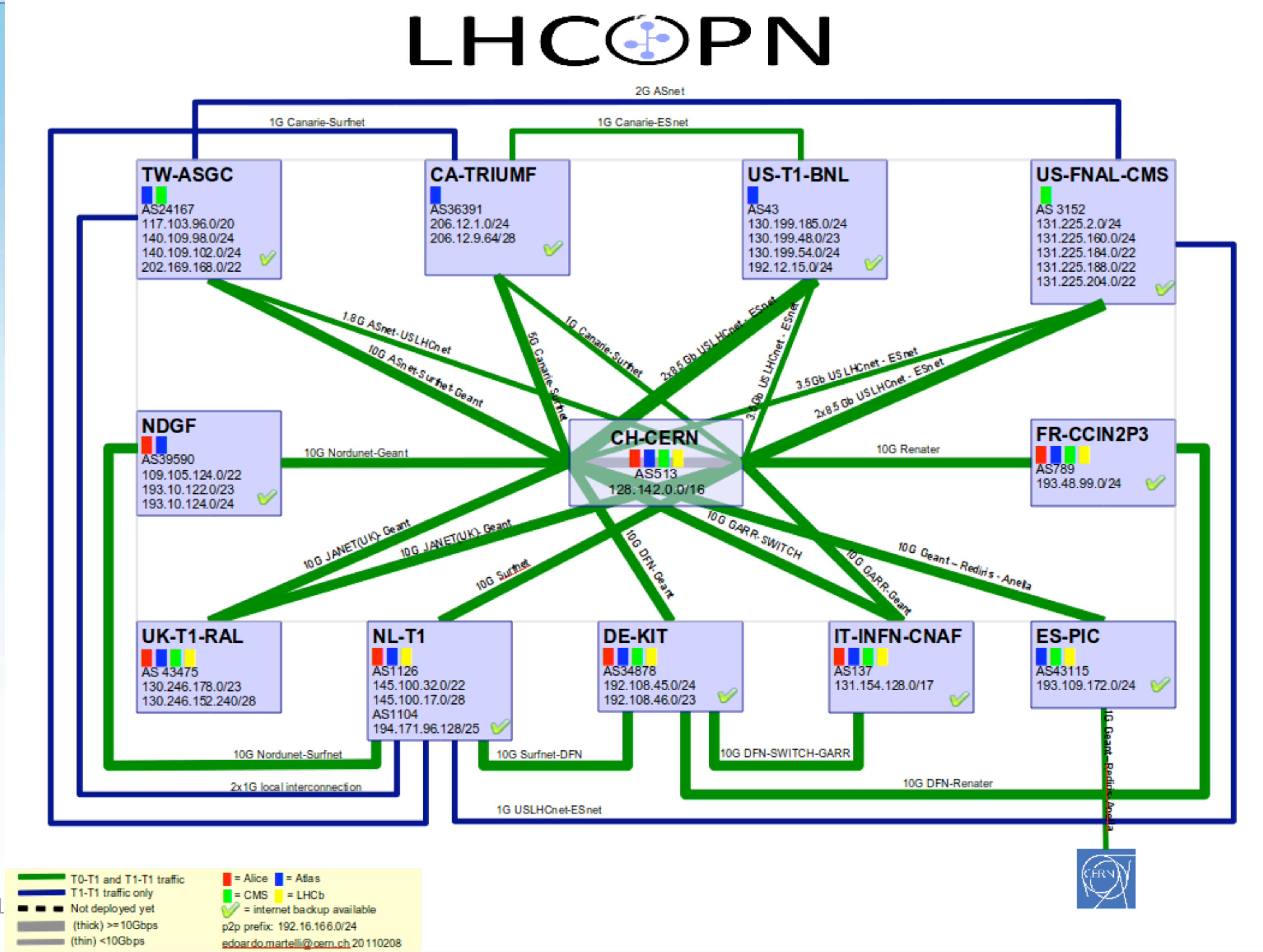


INTERNET2

LHC Traditional Model - LHCOPN

- The LHCOPN is the private IP network that connects the Tier0 and the Tier1 sites of the LCG.
- The LHCOPN consists of any T0-T1 or T1-T1 link which is dedicated to the transport of WLCG traffic and whose utilization is restricted to the Tier0 and the Tier1s.
- Any other T0-T1 or T1-T1 link not dedicated to WLCG traffic may be part of the LHCOPN, assuming the exception is communicated to and agreed by the LHCOPN community

LHC Traditional Model

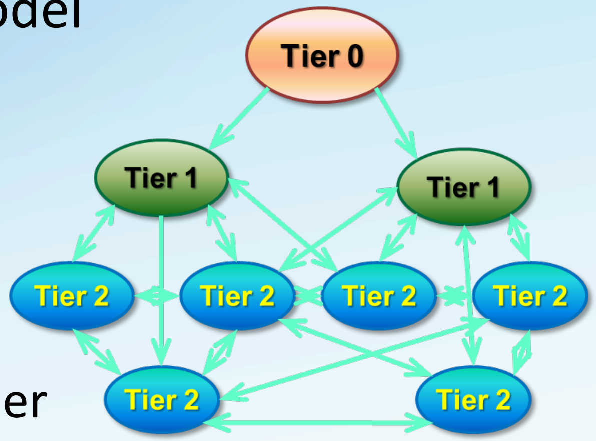


LHC Traditional Model - Tradeoffs

- LHCOPN handled the top layer, what about the others?
- Well stated data movement paths were a positive thing for networking community
  - Predictability for paths to support
  - Sizes were well stated, even though they continued to grow
  - Traffic engineering (at the campus, regional, and backbone layer could support this)
- (Minor) Complications
  - CMS ‘any to any’ model makes engineers nervous
  - Ex: US Sites tracking long sustained flows to sites in the EU – this puts a strain on several portions of the R&E infrastructure
    - Transatlantic Capacity (Limited)
    - Backbone networks
    - Regional networks
    - Campus

New Data Models

- Moving away from the strict MONARC model
- 3 recurring themes:
  - Flat(ter) hierarchy: Any site can use any other site as source of data
  - Dynamic data caching: Analysis sites will pull datasets from other sites "on demand", including from Tier2s in other regions
    - Possibly in combination with strategic pre-placement of data sets
  - Remote data access: jobs executing locally, using data cached at a remote site in quasi-real time
    - Possibly in combination with local caching
- Expect variations by experiment



Source: A. Barczyk

A New Way Forward

- How to support this?
  - This does put a strain on networking support at all levels
  - Architectures, as well as costs
- Changes to infrastructure?
  - Capacity planning may start years in advance in some cases…
  - Could some form of “overlay” be used
- Success factors?
  - Better than current approach
  - Scalable with future needs

Reality Check

- R&E Networking "works for you"
  - Its our job to support activities like LHC
  - Do we have solutions that will scale?
  - If we don't, should we be inventing new ways forward?
- Things to consider
  - Linking in to the campus and regional networks
  - Backbone capacity (lots of science being done…)
  - Ability to expand
- Starting with the first item …

Agenda

- Internet2 Support of Science & Research
- LHC Computing Changes – R&E Networking Response
- Recent Initiatives of Interest
  - NDDI & OS^3E
  - LHCONE
  - DYNES
- Conclusions

Initiative Roundup

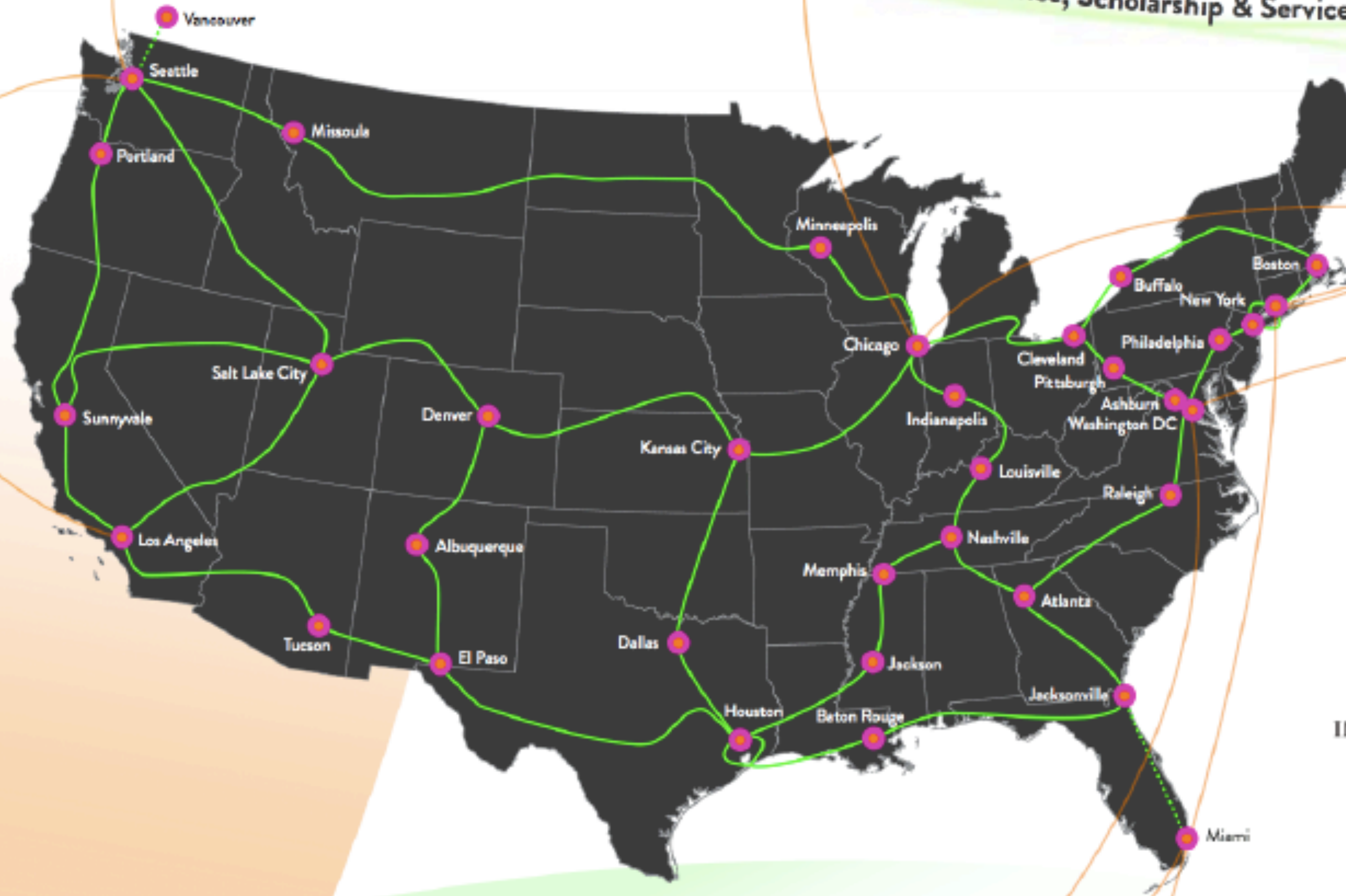
- Initiatives:
  - NDDI and OS^{3}E
    - Partnership to create a new network and software platforms
    - Support global scientific research in a revolutionary new way.
  - LHCONE
    - New network model for supporting the LHC experiments
    - Provide a collection of access locations that are entry points into a network that is private to the LHC T1/2/3 sites.
  - DYNES
    - Nationwide cyber-instrument spanning about 50 Universities and Connectors
    - Supports large, long-distance scientific data flows in the broader scientific community

NDDI and OS^3E

- http://www.internet2.edu/nddi
- Internet2, Indiana University and the Clean Slate Program at Stanford University announced the Network Development and Deployment Initiative ("NDDI")
- New Internet2 service called the Open Science, Scholarship and Services Exchange (OS^3E).
  - OS^3E and NDDI capabilities will be developed and interconnected with links to Europe, Canada, South America and Asia through coordinating international partners like CANARIE in Canada, GÉANT in Europe, JGNX in Japan and RNP in Brazil
  - Additional service partners to be identified.

OS^{3}E

The Open Science, Scholarship & Services Exchange



INTERNET2

INDIANA UNIVERSITY

STANFORD UNIVERSITY

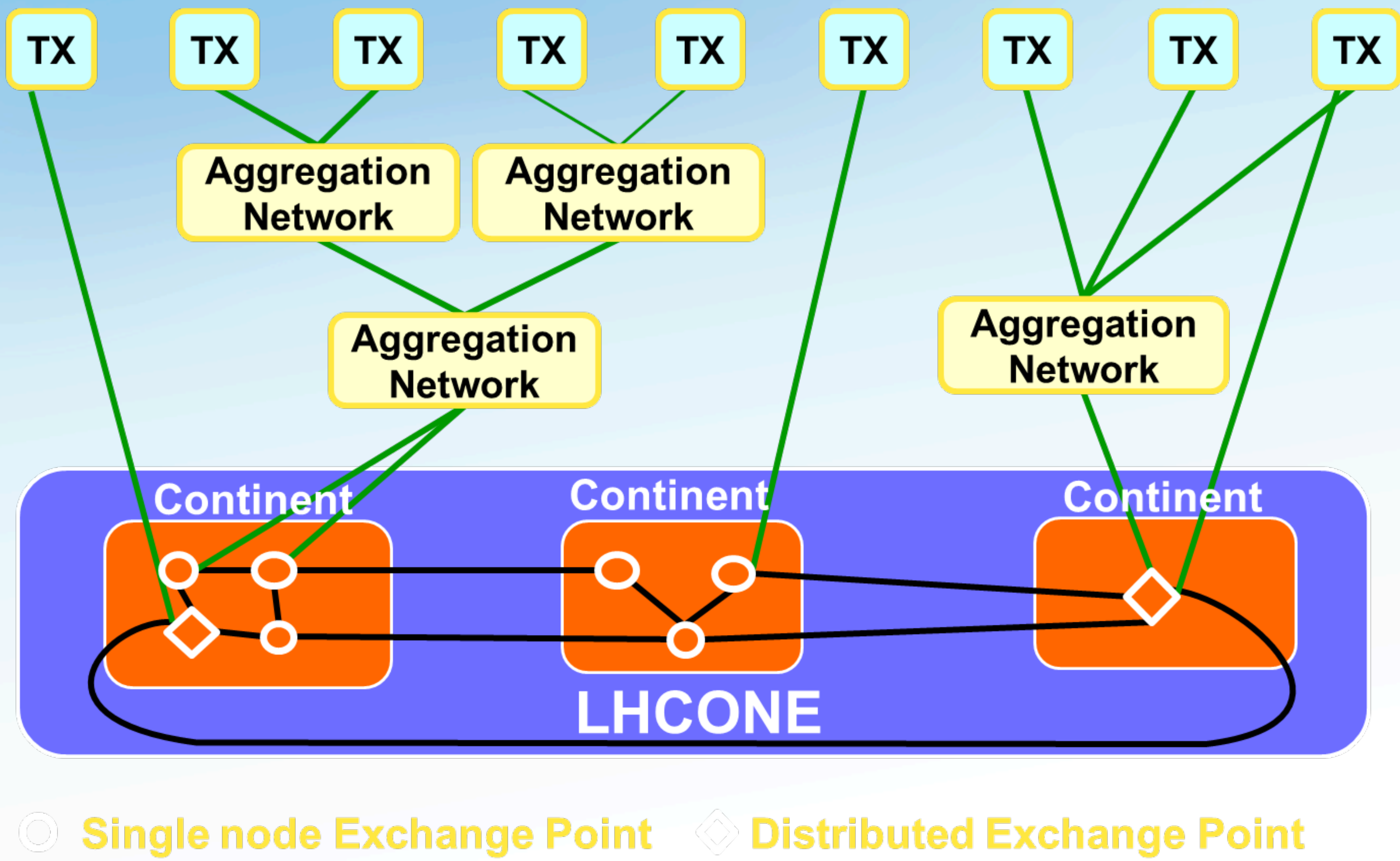
NDDI and OS^3E

- NDDI built using the first nationwide production deployment of OpenFlow technology
  - Will deliver "software defined networking" (SDN) capabilities,
  - Provide a common infrastructure that can create multiple virtual networks, allowing network researchers to test and experiment with new Internet protocols and architectures
  - Enabling domain scientists to accelerate their active research with collaborators worldwide.
- As the first service developed on the NDDI substrate, Internet2 will introduce the Open Science, Scholarship and Services Exchange to offer
  - Persistent VLAN services
  - Nationwide layer 2 open exchange capability

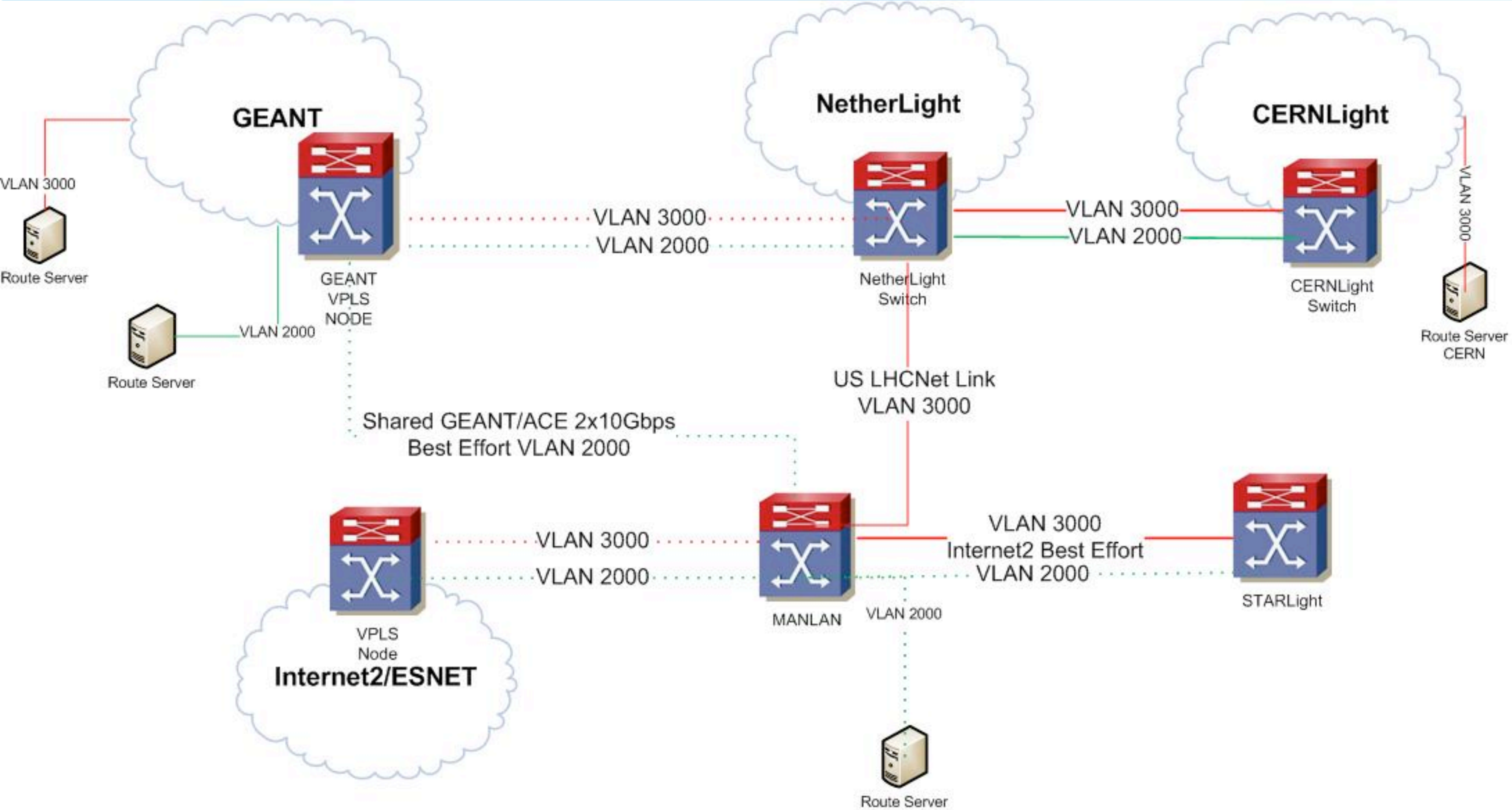
LHCONE Architecture

- Builds on the Hybrid network infrastructures and Open Exchanges
  - To build a global unified service platform for the LHC community
- LHCONE's architecture incorporates the following components
  - Single node Exchange Points
  - Continental / regional Distributed Exchanges
  - Interconnect Circuits between exchange points
    - Likely by allocated bandwidth on various (possibly shared) links to form LHCONE
- Access method to LHCONE is chosen by the end-site, alternatives may include
  - Dynamic circuits
  - Fixed lightpaths
  - Connectivity at Layer 3, where/as appropriate
- Envisioned that many of the Tier-1/2/3s may connect to LHCONE through aggregation networks

LHCONE High-level Architecture



LHCONE – As of this Week …



OS^{3}E & LHCONE

- It is envisioned that future generations of LHCONE (or other privately addressable "VO Overlay Networks)" will ride on top of the OS3E network
  - Ability to control traffic to fine levels
  - Ability to build Layer2 (dynamic or static) and Layer3 connectivity
  - Nationwide availability
  - Integration with existing software frameworks
    - OSCARS for Circuit control
    - perfSONAR-PS for monitoring
    - FDT/GridFTP + Phoebus for data movement)
- Current LHCONE Prototype using dedicated waves on Internet2, USLHCnet, GÉANT

DYNES in Relation to LHCONE

- Solution
  - Dedicated bandwidth (over the entire end to end path) to move scientific data
  - Invoke this "on demand" instead of relying on permanent capacity (cost, complexity)
  - Exists in harmony with traditional IP networking
  - Connect to facilities that scientists need to access
  - Integration with data movement applications
    - Invoke the connectivity when the need it, based on network conditions
- Proposed Deployment:
  - Software and hardware support spanning domain boundaries
    - Campus
    - Regional
    - Backbone
  - Integration with existing technologies and deployments

INTERNET2

DYNES Summary

- What is it?:
  - NSF #0958998
  - A nationwide cyber-instrument spanning ~40 US universities and ~14 Internet2 connectors
    - Extends Internet2's ION service into regional networks and campuses, based on OSCARS implementation of IDC protocol (developed in partnership with ESnet)
    - High-performance file store at sites
- Who is it?
  - A collaborative team including **Internet2**, **Caltech**, **University of Michigan**, and **Vanderbilt University**
  - Community of regional networks and campuses
  - LHC, astrophysics community, OSG, WLCG, other virtual organizations

DYNES Participants

- Application process required to establish participants
  - Submit applications to gauge institutional/network interest
  - Encourage discussion with PIs to advance understanding of the scientific use cases
- Deployment Announcements announced in Feb 2011:
  - 25 End Sites
  - 8 Regional Networks
  - Collaboration with like minded efforts (DoE ESCPS)
- Plans to consider provisional applications (send email to dynes-questions@internet2.edu if interested)

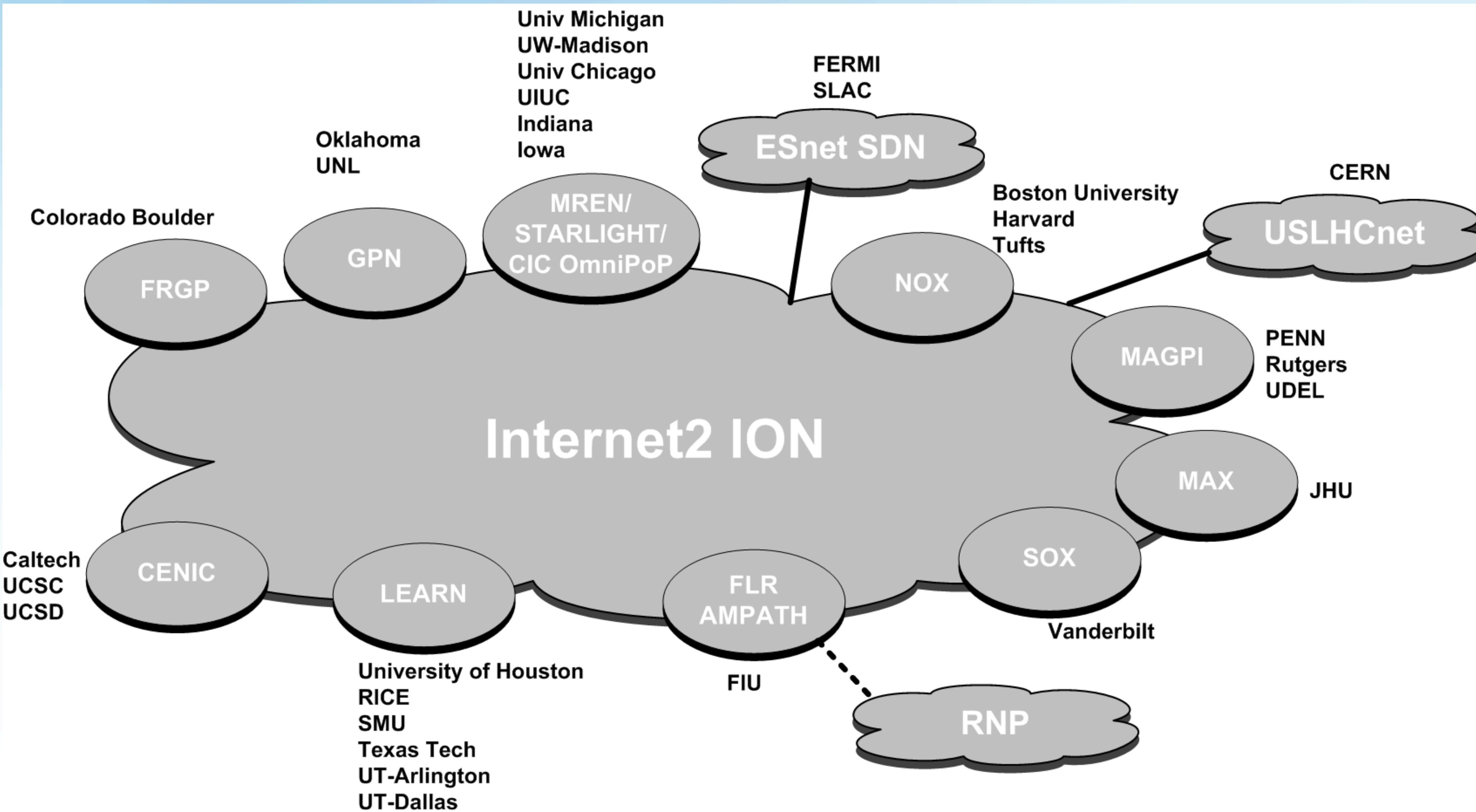
INTERNET2

DYNES Participants

- ***CMS***
  - California Institute of Technology
  - Vanderbilt University
  - Rutgers, The State University of New Jersey
  - The Johns Hopkins University
  - Texas Tech University
  - The University of Nebraska-Lincoln
  - University of Houston and Rice University
  - University of California, San Diego
  - University of Colorado Boulder
  - The University of Iowa
- Pending
  - Florida International University
  - Texas A&M University
  - University of Florida
- ***ATLAS***
  - University of Michigan
  - Indiana University
  - Southern Methodist University (SMU)
  - The University of Chicago
  - The University of Texas at Arlington
  - The University of Pennsylvania
  - Boston University
  - Harvard University
  - The University of Oklahoma
  - University of Illinois at Urbana-Champaign
  - The University of Texas at Dallas
  - Tufts University
  - University of California, Santa Cruz
  - University of Wisconsin
- Pending
  - RENCI/Duke University

DYNES Projected Topology (July 2011)

- Based on applications accepted
- Showing peerings to other Dynamic Circuit Networks (DCN)



DYNES – Why Does this Matter?

- Network Connectivity to some facilities has been problematic
  - Shared link with the rest of campus
  - Smaller R&E link, or Commodity access only
  - Hard to convince CIO level of the need for upgraded capcity
- What does DYNES do?
  - Works with existing R&E connection – no additional connection costs for participants
  - Offers a “private channel” so TCP does not need to compete w/ other campus traffic
  - Connections to all other CMS/ATLAS facilities – through LHCONE the ability to connect anywhere in the world

DYNES Phase 1 Project Schedule

- Phase 1: Site Selection and Planning (Sep-Dec 2010)
  - Applications Due: December 15, 2010
  - Application Reviews: December 15 2010-January 31 2011
  - Participant Selection Announcement: February 1, 2011
- 33 Were Accepted in 2 categories
  - 8 Regional Networks
  - 25 Site Networks

INTERNET2

DYNES Phase 2 Project Schedule

- Phase 2: Initial Development and Deployment (Jan 1-Jun 30, 2011)
  - Initial Site Deployment Complete - February 28, 2011
    - Caltech, Vanderbilt, University of Michigan, MAX, USLHCnet
  - Initial Site Systems Testing and Evaluation (almost) complete: April 29, 2011

- Phase 2 Status:
  - All equipment is in place at initial sites
  - Testing through June 2011

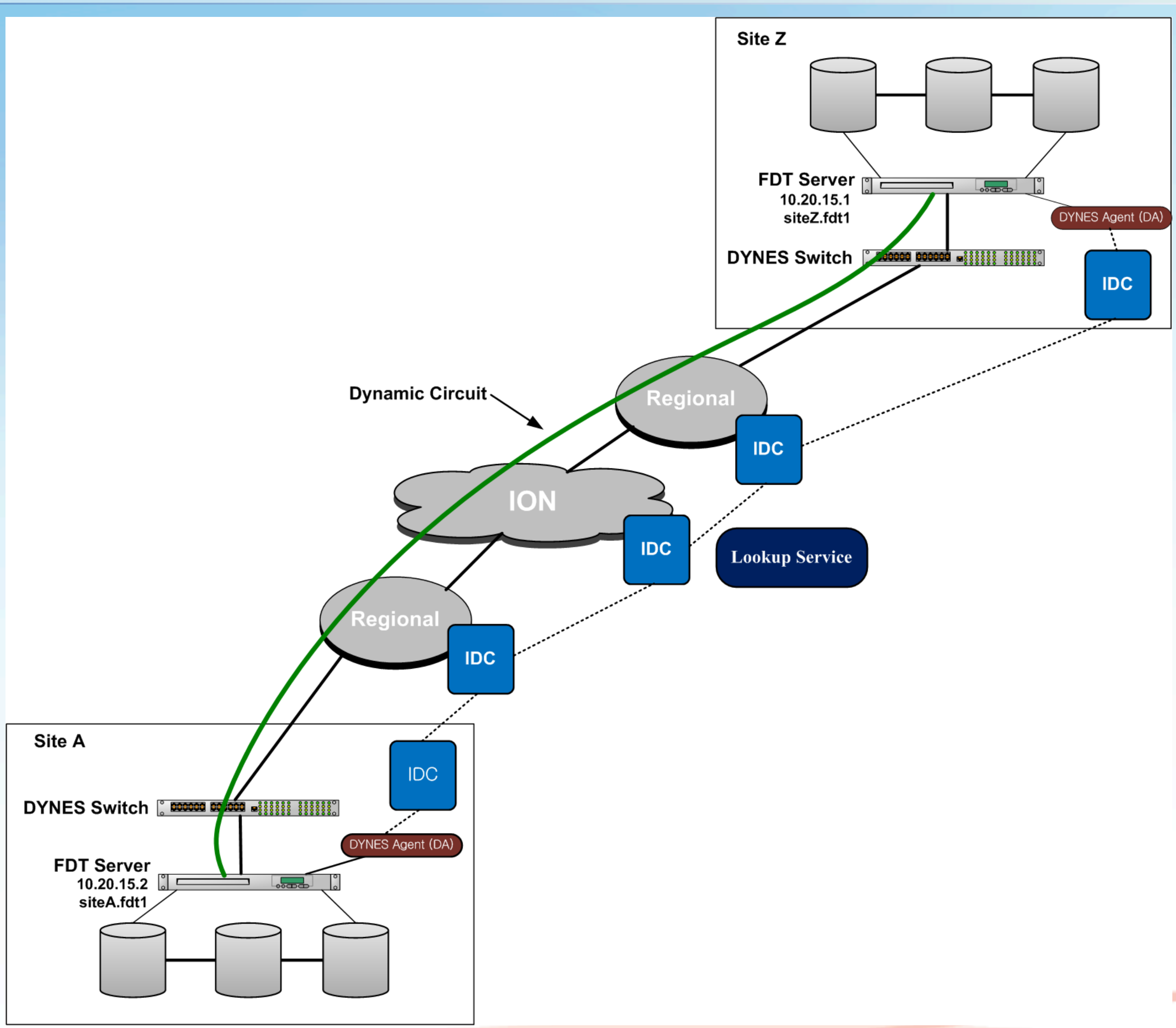
DYNES Phase 3 Project Schedule

- Phase 3: Scale Up to Full-scale System Development (14 months) (July 1, 2011-August 31, 2012)
  - Phase 3-Group A Deployment (9 Sites): March 1-July 1, 2011
  - Phase 3-Group B Deployment (13 Sites): July 18-August 26, 2011
  - Phase 3-Group C Deployment (11 Sites): September 5-October 14, 2011
  - Full-scale System Development, Testing, and Evaluation (October 17 2011- August 31, 2012)
- Phase 4: Full-Scale Integration At-Scale; Transition to Routine O&M (12 months) (September 1, 2012-August 31, 2013)
  - DYNES will be operated, tested, integrated and optimized at scale, transitioning to routine operations and maintenance as soon as this phase is completed

DYNES Software

- Dynamic Circuit Control
  - OSCARS
  - ION Service
- Monitoring
  - perfSONAR Circuit Monitoring
- Data Movement
  - FDT
  - ESCPS

DYNES Data Flow Overview



DYNES Standard Equipment

- Inter-domain Controller (IDC) Server and Software
  - IDC creates virtual LANs (VLANs) dynamically between the FDT server, local campus, and wide area network
  - IDC software is based on the OSCARS and DRAGON software which is packaged together as the DCN Software Suite (DCNSS)
  - DCNSS version correlates to stable tested versions of OSCARS. The current version of DCNSS is v0.5.3.
  - It expected that DCNSSv0.6 will be utilized for Phase 3-Group A deployments and beyond. DCNSSv0.6 will be fully backward compatible with v0.5.3. This will allow us to have a mixed environment as may result depending on actual deployment schedules.
  - The IDC server will be a Dell R410 1U machine.

DYNES Standard Equipment

- Fast Data Transfer (FDT) server
  - Fast Data Transfer (FDT) server connects to the disk array via the SAS controller and runs the FDT software
  - FDT server also hosts the DYNES Agent (DA) Software
  - The standard FDT server will be a DELL 510 server with dual-port Intel X520 DA NIC. This server will a PCIe Gen2.0 card x8 card along with 12 disks for storage.
- DYNES Ethernet switch options:
  - Dell PC6248 (48 1GE ports, 4 10GE capable ports (SFP+, CX4 or optical)
  - Dell PC8024F (24 10GE SFP+ ports, 4 “combo” ports supporting CX4 or optical)

DYNES Equipment and Existing Infrastructure

- DYNES Switch is capable of accepting connections from existing resources at the end site
  - Storage
  - Compute Resources
- Integration of existing resources:
  - Assign ports a "DYNES Friendly" name
  - Use the ION interface to request a circuit between endpoints
  - Use existing software (e.g. GridFTP) to transfer between endpoint locations connected via common DYNES enabled connection

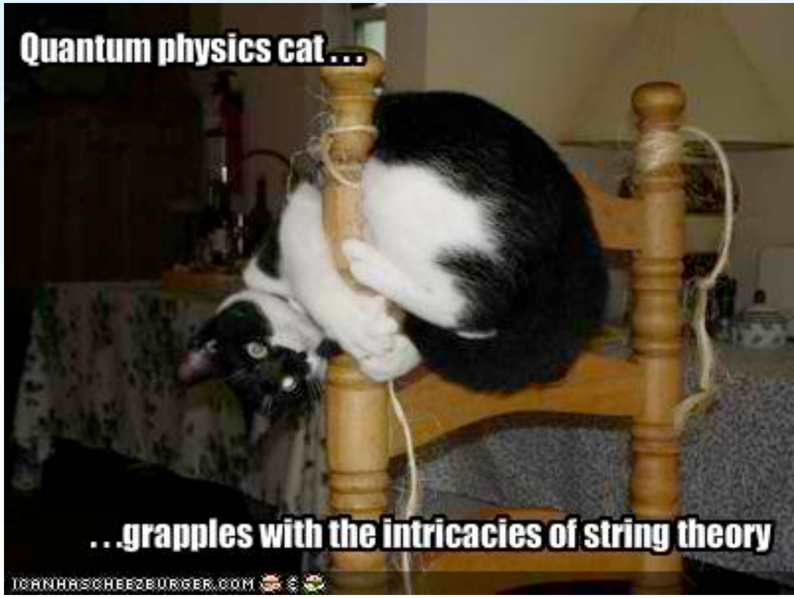
Agenda

- Internet2 Support of Science & Research
- LHC Computing Changes – R&E Networking Response
- Recent Initiatives of Interest
  - NDDI & OS^3E
  - LHCONE
  - DYNES
- Conclusions

INTERNET 2

Conclusions

- Internet2 wants to work closely with domain researchers
  - No matter which discipline
  - We want you to talk with your local IT people too…
- New things happening quickly
  - Demonstrations this fall
  - 100G within the next 2 years
- DYNES still looking for new sites
  - Let us know if you are interested



- zurawski@internet2.edu
- Thanks!

Networking Initiatives for Science

August 11th 2011 – OSG Site Admin Workshop/CMS Tier3 Meeting

Jason Zurawski, Internet2 Research Liaison

For more information, visit https://www.internet2.edu